2055

M.E. Computer Science and Engineering (Cyber Security)

Second Semester

CSN-8205: Information Retrieval

(Common with CS-8304)

Time allowed: 3 Hours   Max. Marks: 50

NOTE: *Attempt <u>five</u> questions in all, including Question No. 1 which is compulsory and selecting two questions from each Unit.*

x-x-x

I.   Attempt the following:-

a) Define tf-idf.

b) What is inverted index?

c) What is probabilistic retrieval model?

d) Explain Relevance Feedback.

e) What is Boolean model?   (5x2)

UNIT - I

II. a) What is tokenization? Explain.

b) Explain what we mean by equivalence classing of terms and why it is useful. Give one
   example of equivalence classing that can fail to retrieve the right documents.   (2x5)

III. a) What is Stemming? Give example.

b) What is query optimization? Explain with suitable example.   (2x5)

IV. What is index construction? Exaplin its various types with example.   (10)

UNIT - II

V.   a) What is the significance of cosine measure in scoring document?

b) What are the components of IR system? Explain each component.   (2x5)

VI. Explain in detail about k-nearest neighbour algorithm.   (10)

VII. Summarize on the working of Web Crawler with its diagram.   (10)

x-x-x